METHODOLOGICAL CONSIDERATIONS IN TOKEN RESEARCH

By

ALAN CONRAD REPP

A Dissertation Presented to the Graduate Council
of the University of Florida
In Partial Fulfillment of the Requirements for
the Degree of Doctor of Philosophy

University of Florida
1973

ACKNOWLEDGEMENTS

This paper acknowledges the amount of teaching by Dr. H. S. Pennypacker and Dr. E. F. Malagodi in Operant Psychology. Further considerations are expressed to Drs. William Reid, William Wolking, James Whorton, and Cary Reichard in Special Education.

For the paper, itself, extreme appreciation is expressed to the teachers who supported this project far beyond the necessity of employer-employee relations: Sylvia Klett, Nancy Speir, Lynn Sosebee, and Cathy Anderson.

Finally, for teaching the friendly union of Skinner and Gibran, I express more than appreciation to Chris.

TABLE OF CONTENTS

		Page
ACKNOWLEDGEMENTS		ii
LIST OF TABLES		v
LIST OF FIGURES		vii
ABSTRACT		viii
CHAPTER		
ONE	INTRODUCTION	1
	Statement of the Purpose	2
TWO	REVIEW OF THE LITERATURE	4
	Introduction	4
	Methodology	4
	Token Paradigm	4
	Research Design	5
	Baseline Procedures	8
	Token Studies	10
	Reinforcing Correct and Punishing Incorrect Responding	10
	Delay of Token Exchange	10
	Summary	11
THREE	DESIGN AND INSTRUMENTATION	13
	Introduction	13
	Experiment I	13
	Subjects	13
	Apparatus	13
	Reinforcers	14
	Procedure	15
	Phase 1	15
	Phase 2	16
	Experiment II	16
	Subjects	16
	Apparatus	16
	Reinforcers	17

CHAPTER Page

Procedure 17
Phase 1 17
Phase 2 18
Summary 19

FOUR RESULTS 19

Introduction 19
Experiment I 20
Experiment II 35
Summary 36
Experiment I 36
Experiment II 50

FIVE DISCUSSION AND CONCLUSIONS 51

Discussion 51
Punishing Incorrect Responding 51
Delayed Token Exchange 52
Baselines 53
Multiple and Concurrent Schedules . . . 54

REFERENCES 56

BIOGRAPHICAL SKETCH 60

LIST OF TABLES

TABLE		Page
1	Mean Correct Responses Per Session in Experiment I.	26
2	Analysis of Variance for Correct Responses in Experiment I. . . .	27
3	Newman-Keuls Test on Differences Between Pairs of Means in Experiment I for Correct Responses.	28
4	Mean Incorrect Responses Per Session in Experiment I.	29
5	Analysis of Variance for Correct Responses in Experiment I. . . .	30
6	Newman-Keuls Test on Differences Between Pairs of Means in Experiment I for Incorrect Responses.	31
7	Mean Choice by Subjects of Each Condition in Experiment I. . . .	32
8	Friedman Analysis on Mean Ranks of Choices of Conditions.	33
9	Newman-Keuls Test on Differences Between Pairs of Means in Experiment I for Choices	34
10	Mean Correct Responses Per Session in Experiment II	41
11	Analysis of Variance for Correct Responses in Experiment II . . .	42
12	Newman-Keuls Test on Differences Between Pairs of Means in Experiment I for Correct Responses.	43

Page

TABLE

13 Mean Incorrect Responses Per Session in Experiment II 44

14 Analysis of Variance for Incorrect Responses in Experiment II . . . 45

15 Mean Choice by Subjects of Each Condition in Experiment II . . . 46

16 Friedman Analysis on Mean Ranks of Choices of Conditions. 47

17 Newman-Keuls Test on Differences Between Pairs of Means in Experiment II. 48

LIST OF FIGURES

FIGURE		Page
1a.	Rate of correct responding in A relative to B (A/A+B) in Experiment I	22
1b.	Rate of correct responding in A relative to B (A/A+B) in Experiment I	23
2a.	Rate of incorrect responding in A relative to B (A/A+B) in Experiment I	24
2b.	Rate of incorrect responding in A relative to B (A/A+B) in Experiment I	25
3a.	Rate of correct responding in A relative to B (A/A+B) in Experiment II.	37
3b.	Rate of correct responding in A relative to B (A/A+B) in Experiment II.	38
4a.	Rate of incorrect responding in A relative to B (A/A+B) in Experiment II.	39
4b.	Rate of incorrect responding in A relative to B (A/A+B) in Experiment II.	40

Abstract of Dissertation Presented to the
Graduate Council of the University of Florida
in Partial Fulfillment of the Requirements for
the Degree of Doctor of Philosophy

METHODOLOGICAL CONSIDERATIONS
IN TOKEN RESEARCH

By

Alan Conrad Repp

March, 1973

Chairman: Dr. William Reid
Major Department: Special Education

Two studies were completed to study the specific effects of punishing errors and of delaying token exchange. In the first experiment, punishing errors reduced incorrect matching responses by mentally retarded students but did not increase correct responding. In the second experiment, delaying token exchange did not decrease correct responding.

In both experiments, reinforcement conditions supported more correct responding than the baseline conditions. In each experiment, a baseline condition of response-produced nonexchangeable tokens supported more correct responding than a no-token baseline condition.

Finally, concurrent schedules (choice of condition by the subject) were compared with multiple schedules (rate of

responding by the subject) in both experiments. Results indicated that concurrent schedules were more sensitive to experimental manipulation than multiple schedules. Each of these four findings was discussed in terms of token programs and general token methodologies.

CHAPTER ONE

INTRODUCTION

Since Staats, Staats, Schultz, and Wolf (1962) and Ayllon and Azrin (1968) began their work with tokens in 1961, increasingly more studies have reported the use of token systems. Although the methods, responses, and populations have varied considerably (see O'Leary and Drabman, 1971; and Axelrod, 1971), these studies have retained the basic components of a token paradigm---a child responds appropriately and is given a token which he exchanges for privileges, snacks, events, etc.

With few exceptions (Wolf, Giles, and Hall, 1968; and Hewett, Taylor, and Artuso, 1969), most studies have used within-subject designs. The most frequently used approach (e.g., O'Leary, Becker, Evans, and Saudargas, 1969; Zimmerman, Zimmerman, and Russell, 1969) has employed a pre-token period (baseline), followed by a token period (experimental), followed by a return-to-baseline period. Another procedure has compared a baseline period in which tokens were administered independently of responding (e.g., five minutes after class began) with an experimental period in which appropriate responses were followed by tokens (Miller and Schneider, 1970). A third procedure has compared a

baseline period in which tokens were not exchangeable with a period in which tokens were exchangeable for backup reinforcers (Bushell, Wrobel, and Michaelis, 1968).

Ferster and Skinner (1957) have noted a procedure (multiple schedule) which alternates stimulus-correlated conditions within the same session rather than between sessions as in the ABA designs previously discussed. This method arranges control both within a single subject and within a single session. Although it is a very popular procedure in the analysis of behavior (see Journal of Experimental Analysis of Behavior), few token studies have employed multiple schedules.

While the multiple-schedule design is often more sensitive to changes than the ABA design, some authors (Repp and Wolking, 1973c; Catania, 1966; Herrnstein, 1970) have discussed its occasional insensitivity to schedule differences. By providing a second type of response, these studies have allowed the subject to control which condition of the multiple schedule is in effect (concurrent schedule) and have found choice to be very sensitive to experimental manipulations.

Statement of the Purpose

The present paper, through two experiments, tests the sensitivity of concurrent schedules to manipulation in two classrooms of retarded children. In the first study, the effects of punishing incorrect responding are assessed;

while in the second study, the effects of delaying reinforcement are assessed. In both studies, two types of baselines (no-token and nonexchangeable-token) are compared and overall analyses of response rates and choice behaviors are made.

CHAPTER TWO

REVIEW OF THE LITERATURE

Introduction

Although a large number of token studies have been done in laboratories (e.g., Wolfe, 1936; Cowles, 1937; Smith, 1939; Kelleher, 1957 a & b, 1958; Malagodi, 1967; Repp and Wolking, 1973 a & b), this discussion is limited to applied behavioral studies with human subjects. This review is concerned, as are the two experiments of this paper, with two aspects of token research: (1) methodological considerations, and (2) the effects of punishing incorrect responses and the effects of delaying reinforcement. Initially, the basic token paradigm and methodology are discussed. Following that presentation, the two specific experimental manipulations are discussed.

Methodology

Token Paradigm

The token paradigm is considered to be a two-component extended chain schedule (Kelleher, 1966). In the first, or production component, tokens are delivered when the subject meets the specified schedule requirement. For example, a token may be delivered after every 50 completed problems, or

perhaps after every minute in which no self-mutilating behavior occurs. After tokens are delivered, the second or exchange component is in effect. In this component, a terminal reinforcer is delivered after the collection of tokens has met its schedule requirement. The subject may be required to collect and then be allowed to exchange 100 tokens for ten minutes of free time, or he may be allowed to exchange tokens only during the third period of the class day. Although the combinations of production and exchange requirements are exhaustive, all token paradigms can be identified by their inclusion of these two components: token production and token exchange.

Research Design

There are two major approaches to experimental designs, and these are classified as being either between-subjects or within-subjects. In the simplest form of the between-subjects design, a manipulation is made to one group (experimental) but not to another (control). Comparisons are then made between the groups. Depending upon the complexity of this design, various criticisms can be made. The most basic criticism, however, remains---individual differences are masked. Researchers using token procedures have usually avoided this approach, and only one study (Hewett, Taylor, and Artuso, 1969) has used a between-subjects design.

The token studies have, instead, used a within-subjects design. The typical design has been to alternate sessions of one condition with sessions of another condition. For

example, Birnbrauer, Wolf, Kidder, and Tague (1965) followed a baseline period in which no tokens were given with a period in which tokens were given. The authors then repeated this alteration. Such a procedure is typically described as an A-B-A-B design although it has been more formally discussed by Sidman (1960) as an intra-subject replication design; by Baer, Wolf, and Risley (1968) as a reversal technique; and by Campbell and Stanley (1963) as an equivalent time-samples design. The popularity of this design has been founded on its reflection of individual differences which are not masked by grouping data from many subjects. The supposition of the A_1-B_1-A_2-B_2 design is that if B_1 differs from A_1 in the same way that B_2 differs from A_2, and if A_2 differs from B_1 in the same way that A_1 differs from B_1, control has been demonstrated.

As with other designs, there are also criticisms of the A-B-A-B procedure. The most frequent criticism is that A_2 is not the same as A_1. The most obvious, though not the only, cause of this difference is that a factor (B_1) has come into contact with the subject's behavior. A second reason is that, in applied studies, the behavior of the technicians of the experiment (usually teachers) is not the same across conditions. Usually (cf. Mandelker, Brigham, and Bushell, 1970), the staff interacts more in the token periods than in the baseline periods. Although the interactions may ensure reversals, the change in behavior across conditions will reflect effects of the teacher's attention

as well as any effects of the token program. In addition, other variables that covary with the token period may be effective. Instructions can be just such a variable (cf. Kazdin and Bootzin, 1972).

Another within-subjects approach is the multiple-baseline design (Baer, Wolf, and Risley, 1968) or multi-element baseline design (Sidman, 1960). In this design, at least two responses are recorded. When the rate of each has stabilized, the token procedure is begun on one of the behaviors. When the response rate of this behavior has stabilized, the procedure is implemented on the other response. In this design, there is no reversal or return-to-baseline condition.

The basis of this design is a multiple schedule (Ferster and Skinner, 1957); that is, a design in which at least two stimulus-correlated conditions alternate within the same session. This design is even more sensitive to individual differences than the ABAB design. Its significant lack of experimental use within token studies can be traced only to the increased difficulty of employing several conditions per day as compared with employing a single condition for several days. As previously mentioned, some authors (Repp and Wolking, 1973c; Catania, 1966; Herrnstein, 1970) have discussed the occasional insensitivity of multiple schedules to schedule differences. The insensitivity of multiple schedules is shared by the ABAB and between-subjects designs as well when they are used in token

studies. The occasional insensitivity is not due to the design, per se, but rather to the datum employed (df. Weiner, 1969, for a discussion of rate's insensitivity to schedule changes). By providing a second type of response, the study by Repp and Wolking (1973c) allowed the subject to control which condition of the multiple schedule was in effect (concurrent schedule) and found choice to be very sensitive to experimental manipulation. The present paper, despite the procedural complexities, employed both multiple and concurrent schedules to assess the effects of the experimental manipulations.

Baseline Procedures

Baseline conditions are those against which the experimental conditions are compared in order to assess the effects of the token program. Within these studies, three types of baselines have been used. Several (e.g., O'Leary, Becker, Evans, and Saudargas, 1969; Zimmerman, Zimmerman, and Russell, 1969; Walker and Buckley, 1968; Birnbrauer, Wolf, Kidder, and Tague, 1965) have employed a no-token phase (baseline), followed by a token (experimental), and then a no-token (return to baseline) phase. These studies have, in effect, compared a phase that included tokens and backup reinforcers with a phase that included neither tokens nor backup reinforcers. Other studies (e.g., Schaefer and Martin, 1966; Miller and Schneider, 1970) have used a baseline period in which tokens were administered independently of responding. These studies have compared a phase that

included response-produced tokens plus backup reinforcers. A third procedure (e.g., Bushell, Wrobel, and Michaelis, 1968) has used a baseline procedure in which tokens were response produced but were not exchangeable. This procedure has compared a phase that included response-produced tokens and no backup reinforcers with a phase that included response-produced tokens and backup reinforcers. Clearly, the third procedure involves response feedback to the subject that is not included in the baseline periods of the first two procedures. If one is assessing the effects of a token program (which is defined as having a token-production and a token-exchange component), different results may occur if these baselines produce different behaviors. If response feedback, as provided when nonexchangeable tokens are response produced, produces more responding than no tokens or response-independent tokens, then experimental phases compared with baseline phases employing the latter two procedures may not be demonstrating only the effectiveness of token programs. The simple administration of a token, indicating to the subject that he is correct, and usually administered with some concomitant attention, may be contributing to the token program's effectiveness without the use of backup reinforcers. The present paper, in both experiments, compared a baseline period in which no token was administered with a baseline period in which nonexchangeable tokens were administered.

Token Studies

Reinforcing Correct and Punishing Incorrect Responding

Experiment I was concerned with the effects of punishing a response within the token paradigm. Most studies within a token system have administered tokens for correct responding and have ignored incorrect responding. Recently, however, several studies have removed tokens for inappropriate responding. Phillips (1968) reported that fines reduced aggressive statements and that the more severe the fine (50 points versus 20 points), the greater the reduction. Others (Siegel, Lenske, and Broen, 1969; Winkler, 1970; and Phillips, Phillips, Fixsen, and Wolf, 1971) reported success in reducing such diverse behaviors as speech disfluencies, noise, violence, and tardiness to meals. Boren and Coleman (1970) found, however, that the addition of a program of point loss for inappropriate behavior to a program of point gain resulted in less appropriate behavior. Although the Boren and Coleman study dealt with two responses (attending meetings and staying in bed), it reported data on only one. The present study simultaneously recorded two responses, correct and incorrect answering, under four conditions: tokens added; tokens added and removed; non-exchangeable tokens; no tokens.

Delay of Token Exchange

Experiment II was concerned with the effects of delaying the exchange of tokens for backup reinforcers. Several token studies (e.g., Phillips, 1968; Zimmerman, Zimmerman,

and Russell, 1969) have reported anecdotally that the interval between token delivery and exchange had been changed throughout the experiment. Atthowe and Krasner (1968) reported a number of different delay periods for several different types of behaviors. O'Leary and Becker (1967) varied the delay between token production and token exchange. Tokens were given for instruction-following behavior and the delay was increased to four days. None of these studies, however, have reported the effects of a delay in exchange on the same response throughout the experiment. The present study compared the effects of a delay in exchange on the subjects' responding.

Summary

Although token studies vary considerably, they share the requirements of the token paradigm: a token-production and a token-exchange component. Virtually all token studies also share the same methodological approach, that being a within-subjects design. Most studies have been of the ABA or ABAB type, in which several baseline sessions are alternated with several treatment sessions. None has attempted to alternate both conditions within the same sessions (multiple schedule). A point of dissimilarity among token studies has been the type of baseline used. Experiments have used baselines in which there were no tokens, response-independent and exchangeable tokens, or response-produced but nonexchangeable tokens.

Additional considerations within specific areas of token research have produced little information in two areas: (a) delaying reinforcement and (b) punishing incorrect responding. While several studies have reported increasing delays of reinforcement, none has provided data on systematic changes in behavior. Studies in punishment, on the other hand, have produced data. The results, however, are confounding. Some report a decrease in behavior when tokens are removed while another indicates an increase.

The present study investigated the use of multiple and concurrent schedules of reinforcement, two baselines, punishing incorrect responding, and reinforcement delay.

CHAPTER THREE

DESIGN AND INSTRUMENTATION

Introduction

Two experiments were conducted investigating variables within the token paradigm. In the first, the effects of punishing incorrect responses were assessed. In the second, the effects of delaying token exchange were assessed. In each experiment, two types of baselines (nonexchange tokens versus no tokens) were used. Data from these baselines were compared with each other as well as with the token conditions. Finally, multiple (rate data) and concurrent (choice data) schedules of reinforcement were compared in each experiment.

Experiment I

Subjects

Seven students in a morning class at the Georgia Retardation Center, Atlanta, Georgia, served as subjects. The class members were ambulatory, between the ages of 11 and 16, and scored between 31 and 55 on the Stanford-Binet Intelligence Test.

Apparatus

The experiment was conducted in a large classroom at the Georgia Retardation Center. At the beginning of the

sessions, each subject was seated at his own desk. On a table in front of the room were four stacks of papers, labeled A, B, C, and D. Each stack consisted of the same material---8-inch by 11-inch (20.3 cm by 27.9 cm) mimeographed pages with various geometric shapes. On each of the pages, there were twenty shapes arranged in five rows and four columns. Each column on the left contained a figure to be matched from one of the three alternatives on the right. The four stacks were differentiated by a different color, two-inch band across the top of each page. Taped to the desk, in front of each stack of paper, was a five-inch by seven-inch (12.7 cm by 17.8 cm) card, colored to match the stack next to it. All stimuli associated with papers A, B, C, and D were colored blue, red, white, and green respectively. For each session, the materials in A, B, C, and D were the same. They did, however, vary between sessions.

Reinforcers

Poker chips, colored blue, red, or white, were used as tokens and were administered for correct responses. Because the ability of each child to do the problems varied, the ratio of tokens to correct responses varied between subjects but, once established for each subject, the ratio remained constant for him across all conditions. In an effort to keep the number of back-up reinforcers relatively constant across subjects, the ratio of tokens to back-up reinforcers varied between subjects but was constant for each subject between conditions. This ratio varied between subjects from

2:1 to 10:1. The reinforcers were various foods chosen by the subject. Available each session were a variety of sugar coated cereals, bite size chocolates, small candy bars, and small cups of various drinks.

Procedure

Phase 1.--During each session, the teacher asked each child to come to the front desk, gave him a paper, and asked the subject what color was at the top of the page. When the subject answered correctly, he returned to his desk. The task, on which each child worked for three weeks prior to the experiment, required a geometric shape to be matched to the sample. At the end of five minutes, each child was asked to stop working. During the five minutes, attendants administered or removed (depending upon the condition) tokens for items completed either correctly or incorrectly. Tokens were exchanged immediately by the subject.

A different condition prevailed for each stack of papers. For stack A, all correct answers were followed by tokens on a 1:1 ratio; all incorrect answers were ignored. On items from stack B, all correct answers resulted in tokens on a 1:1 ratio, and all incorrect answers lost tokens on a 1:1 ratio. A non-exchangeable token was given for each correct answer on papers from stack C while incorrect answers were ignored. No tokens were given for responses to items from stack D.

Approximately three minutes after the five-minute work period ended, another work period began. The same procedure

as before was followed and repeated until each child worked from each stack. The order of each day's presentation (ABCD, BDCA, etc.) was randomized, and a session ended after work was done from each stack. Phase 1 lasted 25 days.

Phase 2.--Phase 2 included the same general procedure as Phase 1, with one major exception. The teacher did not give the papers to the student; rather, the subject chose papers from one stack and worked on the problems for a five-minute period. Again, the subject worked from each of the four stacks. The token conditions correlated with each stack remained the same as they were in the first phase. Phase 2 ended after 20 sessions.

Experiment II

Subjects

Six students in a morning class at the Georgia Retardation Center, Atlanta, Georgia, served as subjects. The class members are ambulatory, between the ages of 13 and 16, and have scored between 39 and 54 on a Stanford-Binet Intelligence test. They were not the same subjects who served in Experiment I.

Apparatus

The research was conducted in a large classroom at the Georgia Retardation Center. Each subject sat at an individual desk. On a table in front of the teacher was four stacks of papers. Each stack consisted of the same material --8-inch by 11-inch (20.3 cm by 27.9 cm) mimeographed pages with various geometric shapes. On each of the pages, twenty

shapes were arranged in five rows and four columns. Each column on the left contained a figure to be matched from one of the three alternatives to the right. Each stack however, had a different color two-inch band across the top of each page. Taped to the desk in front of each stack of papers was a five-inch by seven-inch (12.7 cm by 17.8 cm) card, colored to match the stack above it. All stimuli associated with papers A, B, C, and D were colored blue, red, white, and green. Shapes to be matched remained constant across conditions with each session, but varied between sessions.

Reinforcers

Plker chips, colored blue, red, or white, were used as tokens and were exchanged for backup reinforcers. The ratio of tokens to correct responses varied between subjects but was constant for each subject between conditions. The backup reinforcers were various foods chosen by the child. Available in each session were various packages of sugar coated cereals, bite size chocolates, candy bars, and soft drinks.

Procedure

Phase 1.--During each session, the teacher asked each child to come to the front desk, give him a paper, and asked what color was at the top of the page. When the subject answered correctly, he returned to his chair. The task, on which each child had worked for two weeks prior to the experiment, required a geometric shape to be matched to a sample. During the five minutes, attendants were behind

each child and gave him a token, correlated in color with the band across the page, for the correct responses. Incorrect responses were ignored.

Correct work from stack A was followed by a blue token immediately exchangeable for the subject's choice of food. Work from B was followed by a red token, exchangeable 75 minutes later. A white non-exchangeable token followed work from C, and no token was administered for work from D.

Approximately five minutes after the five-minute work period ended, another work period began. The same procedure as before was followed and was repeated until each child worked from each stack. The session then ended. The duration of Phase 1 was 25 days with the order of presentation (ABCD; DCBA, etc.) randomized.

Phase 2.--Phase 2 included the same general procedure as Phase 1, with one exception. The teacher did not give the papers to the child; rather, the subject chose papers from one stack and worked on the problems for the five-minute period. All token exchange conditions remained the same as in Phase 1. Phase 2 lasted 20 sessions.

Summary

Two studies assessed the effects of punishing errors and of delaying token exchange. Punishment reduced errors and delaying exchange had no effects. The baselines with nonexchangeable tokens produced more responding than the no-token baseline.

CHAPTER FOUR

RESULTS

Introduction

The two experiments presented in this paper were analyzed separately, but similar statistical procedures were used for each. The data in the figures are plotted to represent the rate of correct and incorrect responding in condition A relative to condition B. These data are calculated by dividing the rate in A by the sum of the rates in A and B (A/A+B).

Several types of data were analyzed using various statistical tests, although the same tests were made on the data from each experiment. Within a repeated-measures design (Winer, 1962), an analysis of variance was made on the correct and incorrect responding in each of the four conditions in each experiment. Following that type of analysis, the Newman-Keuls (Winer, 1962) method of testing differences between pairs of means was made for all possible pairs of conditions for both correct and incorrect responding. Overall analysis of choices between conditions was made using Friedman's two-way analysis of variance by ranks (Ferguson, 1959). Following that analysis, the Newman-Keuls test was made on the differences between all possible pairs of means for choices of conditions.

Experiment I

The results of this experiment are shown in Figures 1 and 2 and Tables 1 through 9. Table 1 presents the mean number of correct responses across the four conditions for all sessions, and Table 2 contains the summary of the repeated-measures analysis for this data. The overall analysis indicates that there was a significant difference among responding across conditions, $F_{.99}(3,18) = 28.7$. Following this significant difference, specific tests between conditions were made. In Table 3, the Newman-Keuls test on differences between pairs of means indicates fewer correct responses in D than in C, B, or A; in C than in B or A; but no differences between B and A; $q_{.99}(r,18)$. These results indicate that there was a difference in correct responding between all pairs of conditions except those in which tokens were given for correct responses (A) and tokens were given for correct and taken away for incorrect responses (B).

The data presented in Figures 1a and 1b indicate the rate of correct responding in A relative to B. Each data point was calculated by dividing the rate in A by the sum of the rates in A and B. The relative rate correct for subjects 1-7 was .51, .51, .50, .51, .51, .51, and .52, respectively. Figures 2a and 2b indicate the rate of incorrect responding in A relative to B. These figures were also calculated by dividing the rate in A by the sum of the rates in A and B. The relative rate incorrect for subjects 1-7 was .54, .54, .56, .56, .60, .64, and .55, respectively.

Similar analyses are presented in Tables 4, 5, and 6 for incorrect responses. Table 4 presents the mean number of incorrect responses across the four conditions for all sessions. In Table 5, the repeated-measures test indicates an overall difference between conditions, $F_{.99}(3,18) = 46.00$. The Newman-Keuls test (Table 6) on specific conditions indicates more errors in D than in A, B, or C; in C than in B; and in A than in B; $q_{.99}(r,18)$. The results describe a difference in incorrect responding between all pairs of conditions except those in which exchangeable tokens were given for correct responses (A) and nonexchangeable tokens were given for correct responses (C).

Table 7 presents the mean choices by each subject for each condition. The overall means (with first choice designated as 1; second, as 2; etc.) for conditions A, B, C, and D are 1.46, 2.55, 2.36, and 3.51. Table 8 presents the Friedman analysis of variance by ranks, and the results of this test indicate significance, $\chi^2 = 19.3$. Table 9 presents the Newman-Keuls test on choices and indicates a preference for A to B, C, or D; for C to D; for B to D; but no difference between C and B; $q_{.99}(r,18)$. These results indicate a difference between all pairs of means except that in which nonexchangeable tokens were given for correct responses (C) and that in which tokens were given for correct and removed for incorrect responding (B).

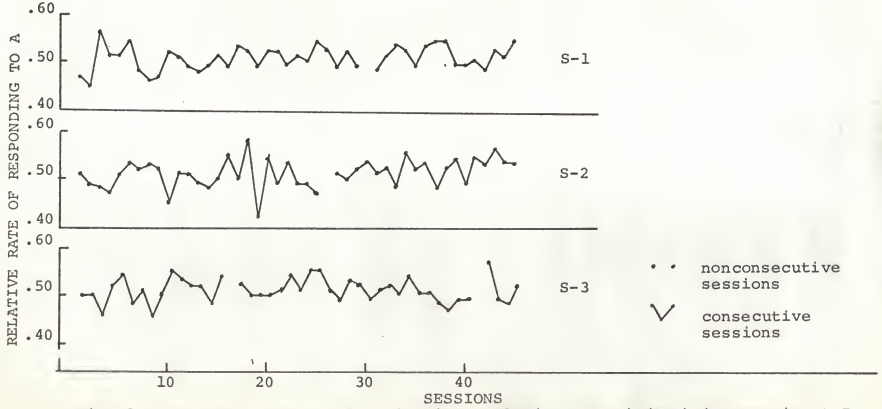

Fig. 1a. Rate of correct responding in A relative to B (A/A+B) in Experiment I.

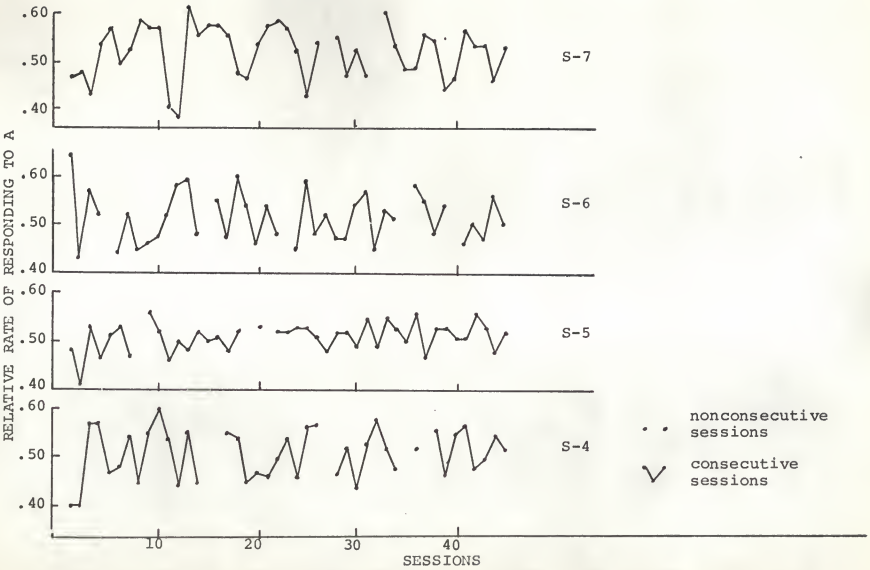

Fig. 1b. Rate of correct responding in A relative to B (A/A+B) in Experiment I.

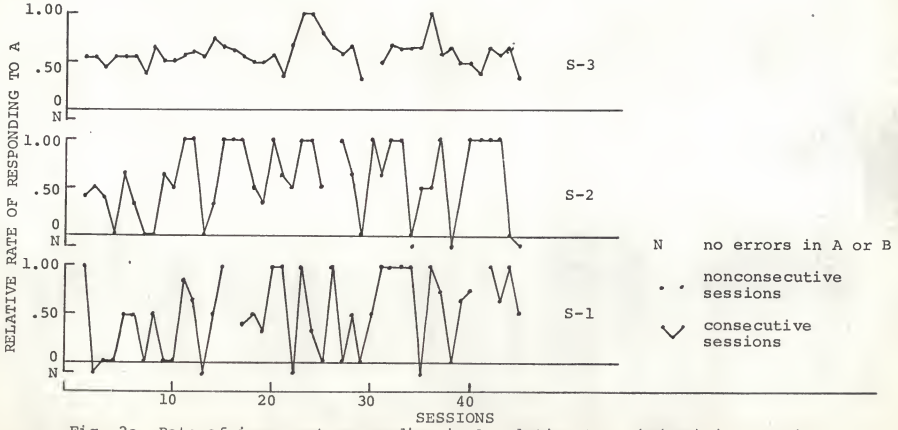

Fig. 2a. Rate of incorrect responding in A relative to B (A/A+B) in Experiment I.

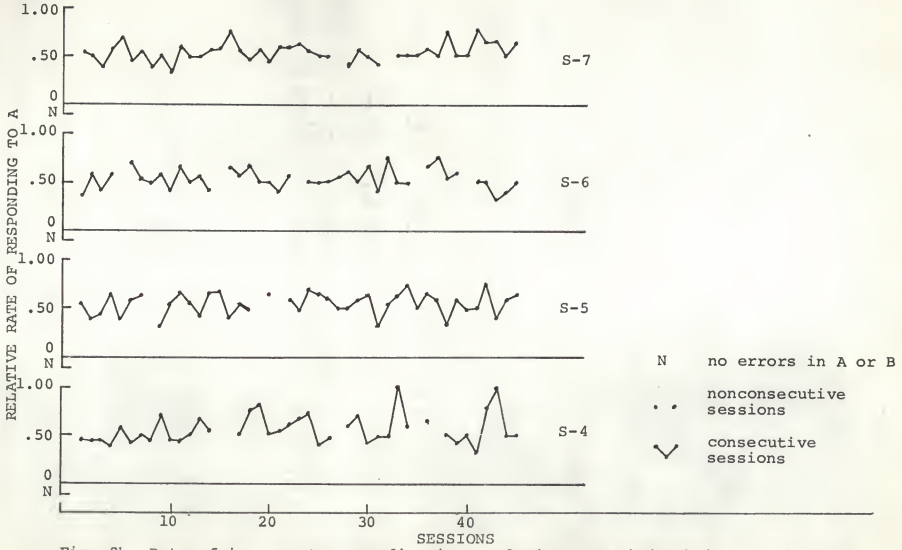

Fig. 2b. Rate of incorrect responding in A relative to B (A/A+B) in Experiment I.

TABLE 1

Mean Correct Responses Per Session in Experiment I

	CONDITION*			
PERSON	A	B	C	D
1	67.84	65.60	62.05	57.40
2	45.53	44.00	41.38	38.80
3	37.59	37.55	34.86	32.95
4	11.48	10.98	10.00	8.95
5	39.26	37.60	35.05	32.05
6	12.70	11.99	10.25	9.08
7	13.90	13.05	10.65	10.14

*A = tokens added
B = tokens added and subtracted
C = nonexchangeable tokens
D = no tokens

TABLE 2

Analysis of Variance for Correct Responses in Experiment I

Source of Variation	SS	df	MS	F
Between people	9600	6		
Within people	156	21		
Treatment	129	3	43	28.67^{xx}
Residual	27	18	1.50	
Total	9756	27		

$^{xx}F_{.99}(3,18) = 5.09$

TABLE 3

Newman-Keuls Test on Differences Between Pairs of Means in Experiment I for Correct Responses

CONDITION		D	C	B	A
	Totals	189	204	221	228
D	189	-	15	32	39
C	204		-	17	24
B	221			-	7
A	228				-
$q_{.99}(r,18)$			4.07	4.70	5.09
$(nMS_{res})^{1/2} = q_{.99}(r,18)(nMS_{res})^{1/2}$			13.19	15.23	16.49

	D	C	B	A
D	-	S	S	S
C		-	S	S
B			-	NS
A				-

TABLE 4

Mean Incorrect Responses Per Session in Experiment I

	CONDITION*			
PERSON	A	B	C	D
1	1.21	0.91	1.42	1.58
2	1.13	0.76	1.31	1.53
3	3.16	2.34	3.57	3.89
4	2.70	2.38	3.45	3.20
5	3.30	2.74	3.36	3.86
6	3.08	2.65	3.15	3.68
7	4.54	4.05	4.84	5.40

*A = tokens added
B = tokens added and subtracted
C = nonexchangeable tokens
D = no tokens

TABLE 5

Analysis of Variance for Correct Responses in Experiment I

Source of Variation	SS	df	MS	F
Between people	36.60	6		
Within people	4.74	21		
Treatment	4.15	3	1.38	46.00^{xx}
Residual	0.59	18	.03	
Total	41.34	27		

$^{xx}F_{.99}(3,18) = 5.09$

TABLE 6

Newman-Keuls Test on Differences Between Pairs of Means in Experiment I for Incorrect Responses

CONDITION		D	C	A	B
	Totals	23.14	21.10	19.12	15.83
D	23.14	-	2.04	4.02	7.31
C	21.10		-	1.98	5.27
A	19.12			-	3.29
B	15.83				-
$q_{.99}(r,18)$			4.07	4.70	5.09
$q_{.99}(r,18)(nMS_{res})^{\frac{1}{2}}$			1.86	2.15	2.33

	D	C	B	A
D	-	S	S	S
C		-	NS	S
A			-	S
B				-

TABLE 7

Mean Choice by Subjects of Each Condition in Experiment I

	CONDITION*			
PERSON	A	B	C	D
1	1.63	2.32	2.53	3.53
2	1.58	2.47	2.05	3.90
3	1.58	2.53	2.37	3.53
4	1.53	2.76	2.24	3.47
5	1.25	2.75	2.35	3.65
6	1.50	2.44	2.45	3.61
7	1.17	2.56	2.50	2.89

*A = tokens added
B = tokens added and subtracted
C = nonexchangeable tokens
D = no tokens

TABLE 8

Friedman Analysis on Mean Ranks of Choices of Conditions by Subjects in Experiment I

	CONDITION			
PERSONS	A	B	C	D
1	1	2	3	4
2	1	3	2	4
3	1	3	2	4
4	1	3	2	4
5	1	3	2	4
6	1	2	3	4
7	1	3	2	4

χ^2 ranks = 19.27

$\chi^2_{.99}(3) = 11.3$

TABLE 9

Newman-Keuls Test on Differences Between Pairs of Means in Experiment I for Choices

CONDITION		A	C	B	D
	Totals	10.24	16.49	17.83	24.58
A	10.24	-	6.25	7.59	14.34
C	16.49		-	1.34	8.09
B	17.83			-	6.75
D	24.58				-
$q_{.99}(r,18)$			4.07	4.70	5.09
$q_{.99}(r,18)(nMS_{res})^{1/2}$			2.65	3.06	3.31

	A	C	B	D
A	-	S	S	S
C		-	NS	S
B			-	S
D				-

Experiment II

The results of this experiment are shown in Figures 3 and 4 and Tables 10 through 17. Table 10 presents the mean number of correct responses in each of the four conditions for all sessions. A summary of the data for the repeated-measures analysis is presented in Table 11. This overall analysis indicates a significant difference among scores across conditions, $F_{.99}(3,15) = 21.2$. Following this significant difference, specific tests between pairs of conditions were made, and an analysis of each of the pairs of means is presented in Table 12. This Newman-Keuls test indicates more correct responding in A than in C or D; in B than in D; in C than in D; but no difference between B-C or A-B; $q_{.99}(r,18)$. These results indicate that there is a difference between all pairs of conditions except two: (1) the condition in which food was delayed 75 minutes (B) and the condition in which nonexchangeable tokens were administered (C); and (2) the condition in which food was delayed 75 minutes (B) and the condition in which food was delivered immediately (A).

The data presented in Figures 3a and 3b indicate the rate of correct responding in A relative to B. The data are calculated by dividing the rate in A by the sum of the rates in A and B. The relative rate correct for subjects 8-13 was .53, .52, .57, .53, .51, and .54, respectively. Figures 4a and 4b indicate the rate of incorrect responding in A relative to B. The figures were also calculated by dividing

the rate in A by the sum of the rates in A and B. The relative rate incorrect for subjects 8-13 was .51, .47, .52, .41, .54, and .41, respectively.

Similar analyses are presented in Tables 13 and 14 for incorrect responding. Table 13 presents the mean number of incorrect responses across the four conditions for all sessions in Experiment II. These data indicate that more errors were made in D than in C, in C than in B, and in B than in A. Table 14 presents the repeated-measures analysis of variance and indicates no significant difference across conditions. As such, the Newman-Keuls test was not made.

Table 15 presents the mean choices by each subject for each condition. The overall means (with the first choice designated as 1; the second, as 2; etc.) for conditions A, B, C, and D are 1.29, 2.22, 2.57, and 3.92, respectively. Table 16 presents the Friedman analysis of variance by ranks, with the results indicating significance, $\chi^2_{.99}(3) = 11.3$. Table 17 presents the Newman-Keuls test on choices and indicates a significant difference between all possible pairs of combinations. The preference between conditions was A>B, A>C, A>D, B>C, B>D, and C>D.

Summary

Experiment I

The overall analysis of the rate data (which is data derived from the multiple schedule) on correct responding indicated differences between conditions. Specific analyses

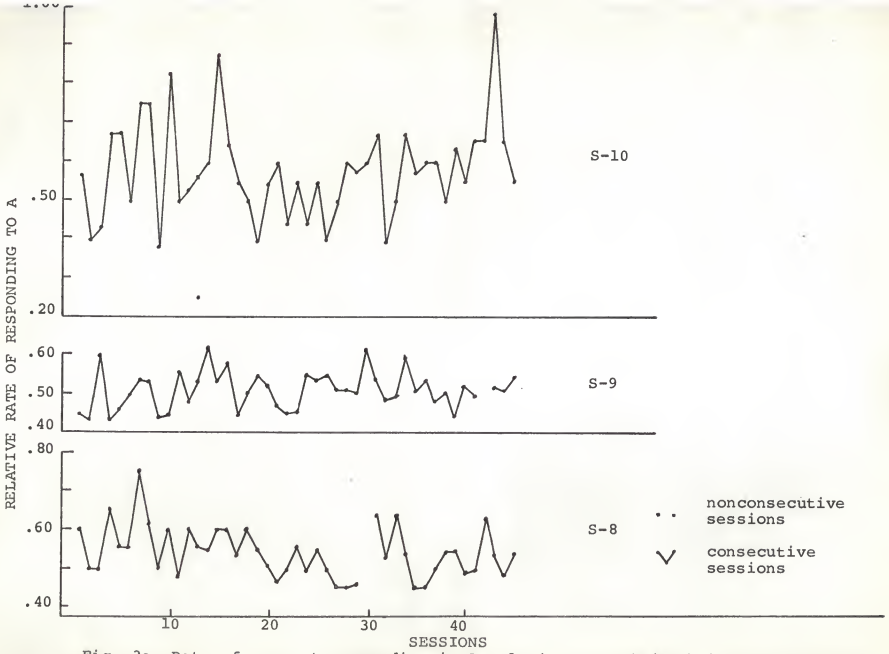

Fig. 3a. Rate of correct responding in A relative to B (A/A+B) in Experiment II.

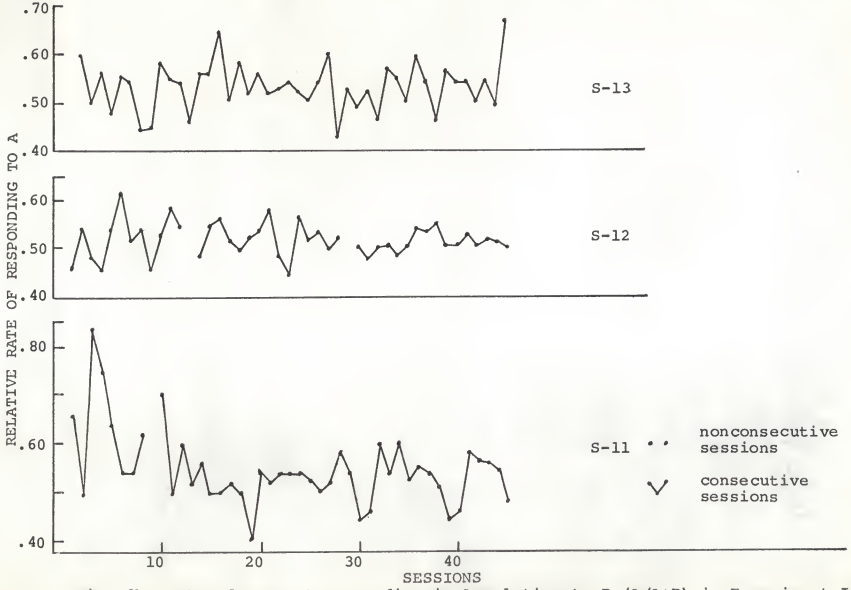

Fig. 3b. Rate of correct responding in A relative to B (A/A+B) in Experiment II.

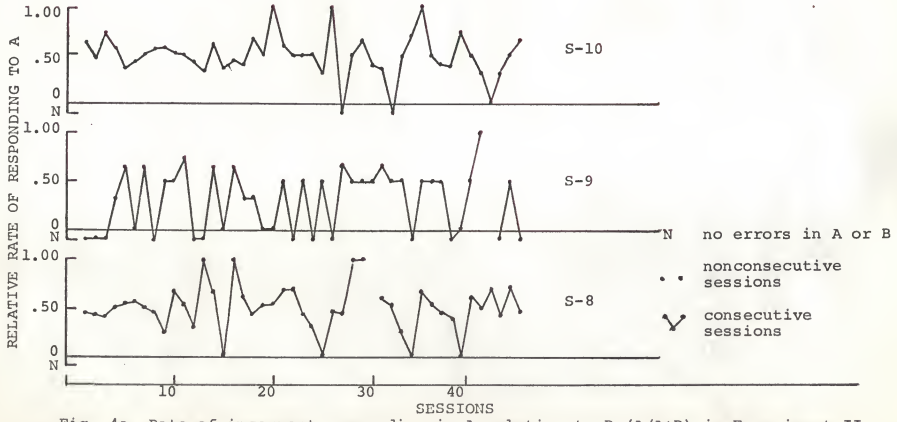

Fig. 4a. Rate of incorrect responding in A relative to B (A/A+B) in Experiment II.

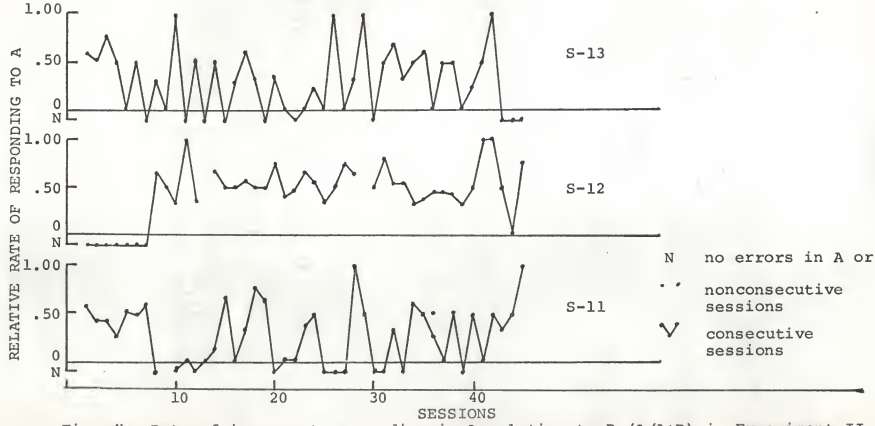

Fig. 4b. Rate of incorrect responding in A relative to B (A/A+B) in Experiment II.

TABLE 10

Mean Correct Responses Per Session in Experiment II

	CONDITION*			
PERSON	A	B	C	D
1	34.77	31.55	27.80	24.27
2	36.41	34.16	31.41	28.55
3	4.20	3.36	3.22	2.71
4	14.23	12.68	11.64	9.93
5	60.35	56.42	54.83	50.83
6	21.09	18.42	17.84	15.00

*A = immediate exchange
B = delayed exchange
C = nonexchangeable tokens
D = no tokens

TABLE 11

Analysis of Variance for Correct Responses in Experiment II

Source of Variation	SS	df	MS	F
Between people	6789	5		
Within people	173	18		
Treatment	140	3	46.67	21.21^{xx}
Residual	33	15	2.20	
Total	6962	23		

$^{xx}F_{.99}(3,15) = 5.42$

TABLE 12

Newman-Keuls Test on Differences Between Pairs of Means in Experiment I for Correct Responses

CONDITION		D	C	B	A
	Totals	131.29	146.74	156.59	171.06
D	131.29	-	15.45	25,30	39.77
C	146.74		-	9.85	24.32
B	156.59			-	14.47
A	171.06				-
$q_{.99}(r,15)$			4.17	4.83	5.25
$q_{.99}(r,15)(nMS_{res})^{1/2}$			15.18	17.58	19.11

	D	C	B	A
D	-	S	S	S
C		-	NS	S
B			-	NS
A				-

TABLE 13

Mean Incorrect Responses Per Session in Experiment II

	CONDITION*			
PERSON	A	B	C	D
1	5.95	5.68	6.59	6.52
2	0.86	0.82	0.98	0.86
3	4.53	4.62	5.04	4.96
4	1.57	1.93	1.84	2.14
5	3.42	3.42	4.74	4.47
6	2.08	3.32	2.44	3.12

*A = immediate exchange
B = delayed exchange
C = nonexchangeable tokens
D = no tokens

TABLE 14

Analysis of Variance for Incorrect Responses in Experiment II

Source of Variation	SS	df	MS	F
Between people	76.74	5		
Within people	3.40	18		
Treatment	1.44	3	.48	3.69^{xx}
Residual	1.96	15	.13	
Total	81.14	23		

$^{xx}F_{.99}(3,15) = 5.42$

TABLE 15

Mean Choice by Subjects of Each Condition in Experiment II

	CONDITION*			
PERSON	A	B	C	D
1	1.53	2.16	2.37	3.95
2	1.16	2.26	2.58	4.00
3	1.05	2.35	2.65	3.95
4	1.35	2.15	2.60	3.90
5	1.21	2.32	2.52	3.95
6	1.45	2.10	2.70	3.75

*A = immediate exchange
B = delayed exchange
C = nonexchangeable tokens
D = no tokens

TABLE 16

Friedman Analysis on Mean Ranks of Choices of Conditions by Subjects in Experiment II

	CONDITION			
PERSON	A	B	C	D
1	1	2	3	4
2	1	2	3	4
3	1	2	3	4
4	1	2	3	4
5	1	2	3	4
6	1	2	3	4

χ^2 ranks = 18.00

$\chi^2_{.99}(3) = 11.3$

TABLE 17

Newman-Keuls Test on Differences Between Pairs of Means in Experiment II for Choices

CONDITION		A	B	C	D
	Totals	7.75	13.34	15.42	23.50
A	7.75	-	5.59	7.67	15.75
B	13.34		-	2.08	10.16
C	15.42			-	8.08
D	23.50				-
$q_{.99}(r,15)$			4.17	4.83	5.25
$q_{.99}(r,15)(nMS_{res})^{1/2}$			1.44	1.67	1.81

	A	B	C	D
A	-	S	S	S
B		-	S	S
C			-	S
D				-

on each of the possible pairs of conditions indicate that for correct responding: (1) the baseline that included non-exchangeable tokens (C) produced more correct responding than the baseline in which there were no tokens (D); (2) that the condition in which correct responses produced exchangeable tokens (A) did not differ from the condition in which correct responses produced tokens and incorrect responses removed tokens (B); and (3) each of the reinforcement conditions (A and B) produced more correct responding than each of the baseline conditions (C and D).

Analysis of incorrect responding indicated: (1) more errors in the no-token baseline condition (D) than in any other; (2) more errors in the other two conditions (C or A) than in the one in which incorrect responses were punished (B); and (3) no difference in errors between C and A; $q_{.99}(r,18)$.

The overall analysis of choice data (derived from the concurrent schedule) indicated significant differences among conditions. Analyses of each possible pair of conditions indicated that: (1) the subjects preferred the condition in which correct responses produced exchangeable tokens (A) to any other condition; (2) they preferred the other two conditions (B and C) to the condition in which there were no tokens (D); and (3) the subjects did not differ in preference between the condition in which tokens were produced for correct and removed for incorrect responses (B) and the condition in which nonexchangeable tokens were produced (C).

Experiment II

These data indicate that delayed exchange (condition B) did not produce less correct responding than immediate exchange (A) or nonexchangeable tokens (C). However, immediate exchange (A) produced more correct responses than either baseline C (no-exchange) or D (no-token); delayed exchange (B) produced more correct responses than no-tokens (D); and no-exchange (C) produced more correct responding than no-tokens (D). Analysis indicated no difference across condition for incorrect responding.

Choice data showed differences not indicated by the rate data. Subjects preferred the immediate exchange condition (A) to all others, the delayed exchange (B) to the two baselines (C,D) and the nonexchangeable-token condition (C) to the no-token condition (D).

CHAPTER FIVE

DISCUSSION AND CONCLUSIONS

Discussion

This paper is discussed in four sections, corresponding to the major considerations: (1) punishing incorrect responding, (2) delayed token exchange, (3) baselines, and (4) multiple and concurrent schedules of reinforcement.

Punishing Incorrect Responding

Among other considerations, Experiment I compared a condition in which correct responses produced tokens and incorrect responses had no effect (A) with a condition in which correct responses produced tokens and incorrect responses lost tokens (B). Statistical analysis indicated that there was no significant differences in correct responding between these two conditions. Inspection of the graphs representing the relative rate of correct responding corroborates the statistical finding.

Data on incorrect responding indicate, however, that more errors were made in A than in B. Both the graphs representing the relative rate incorrect and the statistical analysis indicate that punishing errors by removing tokens reduced their occurrence. This result supports the data of Phillips (1968), Winkler (1970), and Siegel, Lenske, and

Broen (1969), but contradicts the data of Boren and Coleman (1970) who found that removing tokens increased inappropriate responding. There is, of course, a possibility that subject selection produced this disparity in results. In the Boren and Coleman study, soldiers in a psychiatric ward were the subjects. Social pressures may have been evident in this environment. Interestingly, the decrease in incorrect responding had no effect on a related response, correct responding. A supposition might have been that if the incorrect responding decreased, correct responding would have increased. That this relation did not occur supports the point that although classifications of responses might be related (e.g., in this instance a response had to be either correct or incorrect), the rate of occurrence is environmentally controlled.

Delayed Token Exchange

Experiment II compared a condition in which tokens were immediately exchangeable (A) with a condition in which token exchange was delayed 75 minutes. The relative rate correct for each subject was greater in condition A than in condition B. Statistical analysis, however, was not significant. Because each of these conditions was statistically different from each of the baseline conditions, one cannot conclude that delayed reinforcement has no effect, but rather, in this case, that delaying the terminal reinforcer (food) had no effect. The delivery of a conditioned

reinforcer (the token) did have an effect. The lack of difference could be attributable to a fact that delaying the terminal reinforcer for 75 minutes, if the conditioned reinforcer is immediate, simply has no effect. Or this result may be an artifact of the design. Response generalization might have occurred. That is, the response, and its rate, in condition A might have generalized to condition B. If the study had been conducted in the typical ABAB design, with the first condition being the delayed one, a difference might have been found. Whether the lack of difference is due to the power of token delivery or to the design employed in this study is, of course, an empirical question that might be answered by another study.

Baselines

In each experiment, comparisons were made between the rates of responding in two baseline conditions. In one condition (C), nonexchangeable tokens were delivered for correct responses, while in another condition (D), no tokens were present. In each experiment, statistical analysis indicated more correct responding in condition C than in D. In Experiment I, there were statistically fewer errors in condition C than D. This difference was not, however, found in the second experiment. These results indicate that a major difference in reporting results may be caused by the baseline condition an experimenter chooses. Employing a no-token baseline is more likely to produce significant

results than a baseline which consists of response-produced nonexchangeable tokens. Employing the latter baseline would better cancel effects of teacher attention than the former.

Multiple and Concurrent Schedules

The rate data (from the multiple schedule) showed differences between all pairs of conditions except B-A for correct responses and C-A for incorrect responses in Experiment I. The choice data (from the concurrent schedule) showed differences between these conditions but not between C-B. That a difference was shown for the former conditions demonstrates the greater sensitivity of concurrent schedules to experimental manipulation. The failure of this schedule to show a difference in the C-B condition is interesting in reference to an experiment by Repp and Wolking (1973c) in which subjects demonstrated an extreme preference for reinforcement-only to reinforcement-and-punishment although the final probability of reinforcement was equal in both instances. Perhaps, in the first experiment, the power of the token in condition B was being negated by the conflict portion (reinforcement and punishment) of the condition.

In Experiment II, the rate data (from the multiple schedule) indicated differences in correct responding for all but conditions C-B and A-B. For incorrect responding, no difference was found between any of the combinations of conditions. The choice data, however, showed a difference between each of the six possible pairs of conditions. The

immediate-exchange condition was preferred to all others; the delayed-exchange, to the baseline conditions; and the nonexchange baseline to the no-token baseline. Again, greater sensitivity was demonstrated by the concurrent schedule.

The data from each of these experiments suggest the use of concurrent schedules in token studies in an effort to use another datum (choice) to assess variables. This reasoning has also been suggested by Repp and Wolking (1973 a,b) who found: (1) little difference between variable- and fixed-ratio production schedules in terms of rate but large differences in terms of choice, and (2) little difference between magnitude of reinforcement in terms of rate of reinforcement but large differences in terms of choice. The data from these three studies suggest the efficacy of concurrent schedules to other experimenters in token as well as non-token research.

REFERENCES

Atthowe, J. M. and Drasner, L. Preliminary report on the application of contingent reinforcement procedures (token economy) on a "chronic" psychiatric ward. Journal of Abnormal Psychology, 1968, 73, 37-43.

Axelrod, S. Token reinforcement program in special classes. Exceptional Children, 1971, 37, 371-379.

Ayllon, T. and Azrin, N. H. The token economy: A motivational system for therapy and rehabilitation. New York: Appleton-Century-Crofts, 1968.

Baer, D. M., Wolf, M. M., and Risley, T. R. Some current dimensions of applied behavior analysis. Journal of Applied Behavior Analysis, 1968, 1, 91-97.

Birnbrauer, J. S., Wolf, M. M., Kidder, J., and Tague, C. E. Classroom behavior of retarded pupils with token reinforcement. Journal of Experimental Child Psychology, 1965, 2, 219-235.

Boren, J. J. and Coleman, A. S. Some experiments on reinforcement principles within a psychiatric ward for delinquent soldiers. Journal of Applied Behavior Analysis, 1970, 3, 29-37.

Bushell, D., Wrobel, P., and Michaelis, M. Applying "group" contingencies to the classroom study behavior of preschool children. Journal of Applied Behavior Analysis, 1968, 1, 55-62.

Campbell, D. T. and Stanley, J. C. Experimental and quasi-experimental designs for research on training. In W. L. Gage (Ed.), Handbook of research on teaching. Chicago: Rand McNally, 1963. Pp. 171-246.

Catania, A. C. Concurrent Operants. In W. K. Honig (Ed.), Operant behavior: Areas of research and application. New York: Appleton-Century-Crofts, 1966. Pp. 213-270.

Cowles, J. T. Food-tokens as incentives for learning by chimpanzees. Comparative Psychology Monograph, 1937, 14 (5, Serial No. 71).

Ferguson, G. A. Statistical analysis in psychology and education. New York: McGraw-Hill, 1959.

Ferster, C. B. and Skinner, B. F. Schedules of reinforcement. New York: Appleton-Century-Crofts, 1957.

Herrnstein, R. J. On the law of effect. Journal of the Experimental Analysis of Behavior, 1970, 13, 243-266.

Hewett, F. M., Taylor, F. D., and Artuso, A. A. The Santa Monica Project: Evaluation of an engineered classroom design with emotionally disturbed children. Exceptional Children, 1969, 35, 523-529.

Journal of the Experimental Analysis of Behavior. American Psychological Association, 1958.

Kazdin, A. E. and Bootzin, R. R. The token economy: An evaluative review. Journal of Applied Behavior Analysis, 1972, 5, 343-372.

Kelleher, R. T. A multiple schedule of conditioned reinforcement with chimpanzees. Psychological Reports, 1957a, 3, 485-491.

Kelleher, R. T. Conditioned reinforcement in chimpanzees. Journal of Comparative and Physiological Psychology, 1957b, 49, 571-575.

Kelleher, R. T. Fixed-ratio schedules of conditioned reinforcement with chimpanzees. Journal of the Experimental Analysis of Behavior, 1958, 1, 281-289.

Kelleher, R. T. Chaining and conditioned reinforcement. In W. K. Honig (Ed.), Operant behavior: Areas of research and application. New York: Appleton-Century-Crofts, 1966, Pp. 160-212.

Malagodi, E. F. Fixed-ratio schedules of token reinforcement. Psychonomic Science, 1967, 8, 469-470.

Mandelker, A. V., Brigham, T. A., and Bushell, D. The effects of token procedures on a teacher's social contacts with her students. Journal of Applied Behavior Analysis, 1970, 3, 169-174.

Miller, L. K. and Schneider, R. The use of a token system in project Head Start. Journal of Applied Behavior Analysis, 1970, 3, 213-220.

O'Leary, K. D. and Breker, W. C. Behavior modification of an adjustment class: A token reinforcement program. Exceptional Children, 1967, 9, 637-642.

O'Leary, K. D., Becker, W. C., Evans, M. B., and Saudargas, R. A. A token reinforcement program in a public school: A replication and systematic analysis. Journal of Applied Behavior Analysis, 1969, 2, 3-13.

O'Leary, K. D. and Drabman, R. Token reinforcement programs in the classroom: A review. Psychological Bulletin, 1971, 75, 379-398.

Phillips, E. L. Achievement place: Token reinforcement procedures in a home-style rehabilitation center for "predelinquent" boys. Journal of Applied Behavior Analysis, 1968, 1, 213-223.

Phillips, E. L., Phillips, E. A., Fixsen, D. L., and Wolf, M. M. Achievement place: Modification of the behaviors of predelinquent boys within a token economy. Journal of Applied Behavior Analysis, 1971, 4, 45-59.

Repp, A. C. and Wolking, W. D. Preference for variable- versus fixed-ratio production schedules in the token paradigm. Unpublished manuscript, University of Florida, 1973a.

Repp, A. C. and Wolking, W. D. Amount of reinforcement in the token paradigm. Unpublished manuscript, University of Florida, 1973b.

Repp, A. C. and Wolking, W. D. Human preference for non-conflict vs. conflict situations. Psychonomic Science, In press, 1973 c.

Schaefer, H. H. and Martin, P. L. Behavioral therapy for "apathy" of hospitalized schizophrenics. Psychological Report, 1966, 19, 1147-1158.

Siegel, G. M., Lenske, J., and Broen, P. Suppression of normal speech disfluencies through response cost. Journal of Applied Behavior Analysis, 1969, 2, 265-276.

Smith, M. F. The establishment and extinction of the token-reward habit in the cat. Journal of General Psychology, 1939, 20, 475-486.

Staats, A. W., Staats, C. K., Schultz, R. E., and Wolf, M. M. The conditioning of textual responses using "extrinsic" reinforcers. Journal of the Experimental Analysis of Behavior, 1962, 5, 33-40.

Walker, H. and Buckley, N. The use of positive reinforcement in conditioning attending behavior. Journal of Applied Behavior Analysis, 1968, 1, 245-252.

Walker, H. and Buckley, N. The use of positive reinforcement in conditioning attending behavior. Journal of Applied Behavior Analysis, 1968, 1, 245-252.

Weiner, H. Conditioning history and the control of human avoidance and escape responding. Journal of the Experimental Analysis of Behavior, 1969, 12, 1039-1048.

Winer, B. J. Statistical principles in experimental design. New York: McGraw-Hill, 1962.

Winkler, R. C. Management of chronic psychiatric patients by a token reinforcement system. Journal of Applied Behavior Analysis, 1970, 3, 47-55.

Wolf, M. M., Giles, D. K., and Hall, R. V. Experiments with token reinforcement in a remedial classroom. Behavior Research and Therapy, 1968, 8, 51-64.

Wolfe, J. B. Effectiveness of token rewards for chimpanzees. Comparative Psychology Monograph, 1936, 11 (5, Series No. 60).

Zimmerman, E. H., Zimmerman, J., and Russell, C. D. Differential effects of token reinforcement on instruction-following behavior in retarded students instructed as a group. Journal of Applied Behavior Analysis, 1969, 2, 101-112.

BIOGRAPHICAL SKETCH

Alan Conrad Repp was born June 16, 1943, in Miami, Florida. In 1961, he graduated from South Broward High School in Hollywood, Florida. In 1966, he received the degree of Bachelor of Arts, with a major in Psychology, from the University of Florida. In 1970, he received the degree of Master of Arts, with a major in Psychology, from the University of Florida. From that time until present, he has pursued his work toward the degree of Doctor of Philosophy.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

William R. Reid, Chairman
Professor of Special Education

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

James E. Whorton, Co-chairman
Assistant Professor of Special Education

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Cary L. Reichard,
Assistant Professor of Special Education

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

William D. Wolking
Associate Professor of Special Education

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Henry S. Pennypacker
Professor of Psychology

This dissertation was submitted to the Dean of the College of Education and to the Graduate Council, and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

February, 1973

B L Shaw L McBeth

Dean, College of Education

Dean, Graduate School